



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/660,326	09/10/2003	Dinci A. Florencio	MCS-033-03 (304925.01)	5957
27662	7590	06/08/2007	EXAMINER	
MICROSOFT CORPORATION			LERNER, MARTIN	
C/O LYON & HARR, LLP			ART UNIT	PAPER NUMBER
300 ESPLANADE DRIVE			2626	
SUITE 800				
OXNARD, CA 93036				
		MAIL DATE	DELIVERY MODE	
		06/08/2007	PAPER	

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	10/660,326	FLORENCIO ET AL.	
	<b>Examiner</b>	<b>Art Unit</b>	
	Martin Lerner	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on \_\_\_\_\_.
- 2a) This action is FINAL.                    2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1 to 36 is/are pending in the application.
  - 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) Claim(s) \_\_\_\_\_ is/are allowed.
- 6) Claim(s) 1 to 4 and 8 to 36 is/are rejected.
- 7) Claim(s) 5 to 7 is/are objected to.
- 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 10 September 2003 is/are: a) accepted or b) objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    - a) All    b) Some \* c) None of:
      1. Certified copies of the priority documents have been received.
      2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
      3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- 1) Notice of References Cited (PTO-892)
- 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date \_\_\_\_\_.
- 4) Interview Summary (PTO-413)  
 Paper No(s)/Mail Date. \_\_\_\_\_.
- 5) Notice of Informal Patent Application
- 6) Other: \_\_\_\_\_.

## DETAILED ACTION

### *Drawings*

The drawings are objected to because “ $c_i > T$ ” should be “ $E(c_i) > T$ ” in Figure 4, Step 460.

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office Action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as “amended.” If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either “Replacement Sheet” or “New Sheet” pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, Applicants will be notified and informed of any required corrective action in the next Office Action. The objection to the drawings will not be held in abeyance.

***Specification***

The disclosure is objected to because of the following informalities:

On page 7, lines 9 to 13, "TBA" should be updated as "now Application Serial No. 10/663,390 filed 15 September 2003".

On page 18, line 30, "TBA" should be updated as "now Application Serial No. 10/660,325".

On page 28, lines 1 to 6, a reference numeral should be added for computing the energy as "Step 450" of Figure 4.

On page 28, line 19, "may are" should be "may be".

Appropriate correction is required.

***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1 to 4, 10, 15 to 16, 21 to 24, 28 to 32, and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Li et al.* in view of *Clemm*.

(Note: Independent claim 1 is representative of independent claims 1, 15, 24, and 31.)

Concerning independent claims 1, 15, 24, and 31, *Li et al.* discloses a system, computer-implemented process, and method for encoding speech, comprising:

“analyzing sequential segments of at least one digital audio signal to determine segment type as one of speech type segments, non-speech type segments, and unknown type segments” – a comparison of a current frame’s (“sequential segments”) full-band energy to a reference level is made; if the current frame’s energy equals or exceeds the reference level, then a G.729 Annex B VAD (voice activity detector) sets an output to indicate the detected presence of voice activity in the current frame; if the current frame’s energy is less than the reference level, a G.729 Annex B VAD sets its output to zero to indicate the non-detection of voice activity in the current frame (column 9, lines 48 to 63: Figure 3: Steps 23 to 27);

“encoding each speech segment as one or more signal frames using a speech segment-specific encoder” – if VAD 1 detects voice activity, a G.729 speech encoder 3 is invoked to encode the digital representation of the detected voice signal (column 1, line 63 to column 2, line 2: Figure 1); a G.729 encoder is “a speech segment-specific encoder”;

“encoding each non-speech frame as one or more signal frames using a non-speech segment-specific encoder” – however, if VAD 1 does not detect voice activity, a Discontinuous Transmission/Comfort Noise Generator (noise) encoder 2 is used to code the digital representation of the detected background noise signal (column 1, line 63 to column 2, line 2: Figure 1); a Discontinuous Transmission/Comfort Noise Generator (noise) encoder 2 is “a non-speech segment-specific encoder”.

Concerning independent claims 1, 15, 24, and 31, *Li et al.* further discloses that G.729 Annex B performs a multi-boundary initial G.729 Annex B decision to refine an initial decision to reflect the long-term stationary nature of the voice signal. After the initial VAD decision has been smoothed, a final decision is formed. (Column 2, Lines 30 to 38; Column 10, Lines 24 to 37) Thus, *Li et al.* discloses that whether or not a current frame's full-band energy exceeds the reference value is only a first step in determining voice activity, so that there may be frames, through further refinement of the decision, that are equivalent to "unknown type segments". Figure 3 shows that after VAD voice detection is set to 1 or 0, Figure 4 continues with a multi-boundary initial VAD decision to make a smoothed final VAD decision due to background noise, which may change the value of the VAD decision from 0 to 1. If the running averages of the background noise characteristics and supplemental VAD algorithms have diverged, then the values for these characteristics generated by the supplemental VAD algorithm are substituted for the respective values of these characteristics generated by the G.729 Annex B algorithm. (Column 12, Lines 5 to 12; Figure 4: Steps 30 and 41) Subsequently, either speech encoder 3 or Discontinuous Transmission/Comfort Noise Generator (noise) encoder 2 is used to code the digital representation of the voice signal or background noise signal according to the refinement of the final decision. (Column 1, Line 63 to Column 2, Line 2; Figure 1)

Concerning independent claims 1, 15, 24, and 31, the only element omitted by *Li et al.* is the step of "buffering each sequential unknown type segment in a segment buffer until analysis of a subsequent segment identifies the subsequent segment type

as any of a speech segment and silence segment". Generally, it is known that speech encoding involves buffering during processing, implicitly, but buffering is not expressly disclosed by *Li et al.* *Clemm* teaches VAD-directed silence suppression, where a voice signal is received in a buffer during a delay between a start of voice activity and the detection of voice activity. (Column 1, Lines 28 to 58: Figure 1) An objective of buffering voice signals is to ensure that no voice activity is lost during the period of time necessary to turn off silence suppression. (Column 1, Lines 49 to 54) It would have been obvious to one having ordinary skill in the art to buffer segments as suggested by *Clemm* in a method for encoding voice activity by G.729 Annex B of "unknown type segments" of *Li et al.* for a purpose of ensuring that no voice signals are lost during a period of time to determine whether a speech signal has voice activity or no voice activity.

Concerning claim 2, *Li et al.* discloses detection of voice activity and non-voice activity for background noise (column 1, line 63 to column 2, line 2: Figure 1); initially, a current frame's energy is compared to a reference level to determine whether voice activity is detected (column 9, lines 48 to 63); thus, an initial decision reflects whether the current frame is speech or silence.

Concerning claim 3, *Clemm* teaches that the speed of playback may be increased to 150% speed playback when the buffer is full, according to the depletion level of the buffer (column 3, lines 30 to 42: Figures 3A and 3B); increasing the speed of playback before transmission is equivalent to "a burst transmission at a higher rate than a current sampling rate of the audio signal".

Concerning claims 4 and 22, *Clemm* teaches a depletion device flushes the buffer in an accelerated manner when the VAD function is released (column 4, lines 38 to 39).

Concerning claim 10, *Li et al.* discloses a decoder for receiving voice and noise encoded signals (column 2, lines 7 to 14: Figure 1); *Clemm* teaches “a burst transmission” as a speed of playback may be increased to 150% speed playback when the buffer is full, according to the depletion level of the buffer before transmission by transmission device 450 (column 3, lines 30 to 42: Figures 1, 3A, 3B, and 4); implicitly, *Li et al.* operates at a fixed frame rate.

Concerning claim 16, *Clemm* teaches a voice signal may be condensed (“compressed”) by dropping, or removing, packets from the voice signal (column 2, lines 50 to 58); inter-sound space may be compressed, or packets may be dropped, during a condensed playout period (column 3, lines 43 to 49).

Concerning claims 21, 30, and 36, *Li et al.* discloses that speech encoder 3 and Discontinuous Transmission/Comfort Noise (noise) encoder 2 code digital representations of a detected voice signal and a detected background noise signal, respectively, which are transmitted over a communications channel 4 (column 1, line 63 to column 2, line 2: Figure 1); implicitly, speech encoders process signals quickly enough to substantially operate as “a real-time” communications device.

Concerning claim 23, *Li et al.* discloses speech encoder 3 encodes a detected voice signal and Discontinuous Transmission/Comfort Noise (noise) encoder 2 encodes a detected background noise signal (column 1, line 63 to column 2, line 2: Figure 1);

thus, speech encoder 3 and Discontinuous Transmission/Comfort Noise (noise) encoder 2 are "a frame type-specific encoder corresponding to the type of each frame".

Concerning claim 28, *Clemm* discloses preserving the pitch by only compressing inter-sound space, so that the voice perception is more natural (column 3, lines 43 to 49).

Concerning claim 29, *Clemm* discloses dropping, or removing packets from the signal (column 2, lines 55 to 59), which is equivalent to "decimating at least one of the buffered frames."

Concerning claim 32, *Clemm* discloses "temporally compressing the frame" by increasing the playback speed while compressing inter-sound space (column 3, lines 30 to 49).

Claims 8 to 9, 17 to 20, 25 to 27, and 33 to 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Li et al.* in view of *Clemm* as applied to claims 1, 15, 16, 24, 31, and 32 above, and further in view of *Lakaniemi et al.*

*Li et al.* discloses that an initial VAD decision is made using multi-boundary decision regions, but does not expressly disclose identifying an onset point of speech, where an onset point can be identified and encoded as non-speech segments or as speech segments. However, it is known to classify speech frames in a variety of ways, including as onset frames. *Lakaniemi et al.* teaches classifying speech frames into frame types, with frames having lower priority, such as non-speech frames, being selected for control message data, and frames having higher priority frame types, such

as onset and transient frames, being avoided for selection due to the higher subjective contribution to speech quality. (Abstract) (*Clemm* teaches the features of buffering, temporally compressing, and discarding frames, as noted above.) It would have been obvious to one having ordinary skill in the art to identify an actual onset point of speech in a current segment as taught by *Lakaniemi et al.* in a method for encoding voice activity by G.729 Annex B of *Li et al.* for a purpose of avoiding selecting frames having a higher subjective contribution to speech quality for control message data.

Claims 11 to 14 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Li et al.* in view of *Clemm* as applied to claims 1 and 3 above, and further in view of *Ramjee et al.* ("Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks").

*Li et al.* omits a decoder that uses extra samples contained in a burst transmission to populate a jitter buffer for an adaptive playout scheme, where at least some of the received data is compressed to reduce average signal delay. However, *Ramjee et al.* teaches an adaptive playout mechanism, where received audio packets are buffered, and their playout delayed at the destination host in order to compensate for variable network delays. (Abstract) *Ramjee et al.* discloses a delay jitter (Page 2, Left Column) for a buffer having a maximum size (Page 1, Right Column), which is equivalent to a "jitter buffer". The algorithm is applied to talkspurts, which are equivalent to "burst transmission". It would have been obvious to one having ordinary skill in the art to employ the adaptive playout mechanism with a jitter buffer of *Ramjee et al.*.

in a method for encoding voice activity by G.729 Annex B of *Li et al.* for a purpose of compensating for variable network delays.

***Allowable Subject Matter***

Claims 5 to 7 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

***Conclusion***

The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Leitch et al., Zhang, Kapanen, Malah, Ashley, Nayak, and Fayad et al. disclose related art.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML  
5/31/07

Martin Lerner  
Martin Lerner  
Examiner  
Group Art Unit 2626